Reminder No. 2: Propagation of Error, and Standard Errors for Derived Quantities

36-402, Advanced Data Analysis*

Last update: 27 February 2013

A reminder about how we get approximate standard errors for functions of quantities which are themselves estimated with error.

Suppose we are trying to estimate some quantity θ. We compute an estimate $\widehat{\theta}$, based on our data. Since our data is more or less random, so is $\widehat{\theta}$. One convenient way of measuring the purely statistical noise or uncertainty in $\widehat{\theta}$ is its standard deviation. This is the **standard error** of our estimate of θ.[1] Standard errors are not the only way of summarizing this noise, nor a completely sufficient way, but they are often useful.

Suppose that our estimate $\widehat{\theta}$ is a function of some intermediate quantities $\widehat{\psi_1}, \widehat{\psi_2}, \ldots, \widehat{\psi_p}$, which are also estimated:

$$\widehat{\theta} = f(\widehat{\psi_1}, \widehat{\psi_2}, \ldots \widehat{\psi_p}) \tag{1}$$

For instance, θ might be the difference in expected values between two groups, with ψ_1 and ψ_2 the expected values in the two groups, and $f(\psi_1, \psi_2) = \psi_1 - \psi_2$. If we have a standard error for each of the original quantities $\widehat{\psi_i}$, it would seem like we should be able to get a standard error for the **derived quantity** $\widehat{\theta}$. There is in fact a simple if approximate way of doing so, which is called **propagation of error**[2].

We start with (what else?) a Taylor expansion. We'll write ψ_i^* for the true (ensemble or population) value which is estimated by $\widehat{\psi_i}$.

$$f(\psi_1^*, \psi_2^*, \ldots \psi_p^*) \approx f(\widehat{\psi_1}, \widehat{\psi_2}, \ldots \widehat{\psi_p}) + \sum_{i=1}^{p} (\psi_i^* - \widehat{\psi_i}) \left.\frac{\partial f}{\partial \psi_i}\right|_{\psi=\widehat{\psi}} \tag{2}$$

$$f(\widehat{\psi_1}, \widehat{\psi_2}, \ldots \widehat{\psi_p}) \approx f(\psi_1^*, \psi_2^*, \ldots \psi_p^*) + \sum_{i=1}^{p} (\widehat{\psi_i} - \psi_i^*) \left.\frac{\partial f}{\partial \psi_i}\right|_{\psi=\widehat{\psi}} \tag{3}$$

---

*Thanks to Prof. Howard Seltman for corrections

[1] It is not, of course, to be confused with the standard deviation of the data. It is not even to be confused with the standard error of the mean, unless θ is the expected value of the data and $\widehat{\theta}$ is the sample mean.

[2] Or, sometimes, the **delta method**.

$$\hat{\theta} \approx \theta^* + \sum_{i=1}^{p} (\widehat{\psi_i} - \psi_i^*) f_i'(\widehat{\psi}) \quad (4)$$

introducing f_i' as an abbreviation for $\frac{\partial f}{\partial \psi_i}$. The left-hand side is now the quantity whose standard error we want. I have done this manipulation because now $\hat{\theta}$ is a linear function (approximately!) of some random quantities whose variances we know, and some derivatives which we can calculate.

Remember the rules for arithmetic with variances: if X and Y are random variables, and a, b and c are constants,

$$\begin{aligned} \mathrm{Var}[a] &= 0 & (5)\\ \mathrm{Var}[a + bX] &= b^2 \mathrm{Var}[X] & (6)\\ \mathrm{Var}[a + bX + cY] &= b^2 \mathrm{Var}[X] + c^2 \mathrm{Var}[Y] + 2bc\mathrm{Cov}[X, Y] & (7) \end{aligned}$$

While we don't know $f(\psi_1^*, \psi_2^*, \ldots \psi_p^*)$, it's constant, so it has variance 0. Similarly, $\mathrm{Var}\left[\widehat{\psi_i} - \psi_i^*\right] = \mathrm{Var}\left[\widehat{\psi_i}\right]$. Repeatedly applying these rules to Eq. 4,

$$\mathrm{Var}\left[\widehat{\theta}\right] \approx \sum_{i=1}^{p} (f_i'(\widehat{\psi}))^2 \mathrm{Var}\left[\widehat{\psi_i}\right] + 2\sum_{i=1}^{p-1} \sum_{j=i+1}^{p} f_i'(\widehat{\psi}) f_j'(\widehat{\psi}) \mathrm{Cov}\left[\widehat{\psi_i}, \widehat{\psi_j}\right] \quad (8)$$

The standard error for $\widehat{\theta}$ would then be the square root of this.

If we follow this rule for the simple case of group differences, $f(\psi_1, \psi_2) = \psi_1 - \psi_2$, we find that

$$\mathrm{Var}\left[\widehat{\theta}\right] = \mathrm{Var}\left[\widehat{\psi_1}\right] + \mathrm{Var}\left[\widehat{\psi_2}\right] - 2\mathrm{Cov}\left[\widehat{\psi_1}, \widehat{\psi_2}\right] \quad (9)$$

just as we would find from the basic rules for arithmetic with variances. The approximation in Eq. 8 comes from the nonlinearities in f.

If the estimates of the initial quantities are uncorrelated, Eq. 8 simplifies to

$$\mathrm{Var}\left[\widehat{\theta}\right] \approx \sum_{i=1}^{p} (f_i'(\widehat{\psi}))^2 \mathrm{Var}\left[\widehat{\psi_i}\right] \quad (10)$$

and, again, the standard error of $\widehat{\theta}$ would be the square root of this. The special case of Eq. 10 is sometimes called *the* propagation of error formula, but I think it's better to use that name for the more general Eq. 8.